

# Selfish Cops and Passive Robber: Qualitative Games

Ath. Kehagias and G. Konstantinidis

November 22, 2016

## Abstract

Several variants of the *cops and robbers* (CR) game have been studied in the past. In this paper we examine a novel variant, which is played between two cops, each one *independently* trying to catch a “passive robber”. We will call this the *Selfish Cops and Passive Robber* (SCPR) game. In short, SCPR is a stochastic two-player, zero-sum game where *the opponents are the two cop players*. We study sequential and concurrent versions of the SCPR game. For both cases we prove the existence of value and optimal strategies and present algorithms for the computation of these.

## 1 Introduction

Several variants of the *cops and robbers* (CR) game have been studied in the past. In this paper we examine a novel variant, which is played between two cops, each one *independently* trying to catch a “passive robber”. We will call this the *Selfish Cops and Passive Robber* (SCPR) game. Here is a brief and informal description of the game; a more detailed description will be provided in later sections.

1. The game is played on an undirected, finite, simple and connected graph.
2. The game is played by two cop players  $C_1$  and  $C_2$ , each controlling a *cop token* (the tokens will also be referred to as  $C_1$  and  $C_2$ ).
3. A *robber token*  $R$  is also used, which is controlled by Chance.
4. At every turn of the game the tokens are moved from vertex to vertex, along the edges of the graph.
5. The winner is the first player whose token lies at the same vertex as the robber token (i.e., the player who “captures the robber”).

In short, SCPR is a game where *the opponents are the two cop players*. As far as we know this CR variant has not been previously studied.

The study of Cops and Robbers was initiated by Quillot [21] and Nowakowski and Winkler [19]. For the graph theoretic point of the view, the reader can consult the recent book [20] which contains a good overview of the extensive literature. We will study SCPR from a somewhat different angle, using the theory of stochastic games<sup>1</sup> as presented in, e.g., the book by Filar and Vrieze [13]. In the current paper we study *qualitative* SCPR games, i.e., games in which the *payoff* is the *winning probability*; in a forthcoming paper we will discuss *quantitative* SCPR games, i.e., games in which the payoff is the *expected capture time*.

The paper is organized as follows. In Section 2 we present preliminary definitions and notation. In Section 3 we study the *sequential* version of the SCPR game and in Section 4 the *concurrent* version. In Section 5 we discuss connections to several other research areas (for instance *recursive games*, *graphical games* and *reachability games*) and in Section 6 we present concluding remarks and discuss future research directions.

---

<sup>1</sup>I.e., a sequence of normal-form games where the game played at any time depends probabilistically on the previous game played and the actions of the agents in that game.

## 2 Preliminaries

The SCPR game is played on an undirected, finite, simple and connected graph  $G = (V, E)$ , where  $V$  is the vertex set and  $E$  is the edge set. Unless otherwise stated, we will assume that the *cop number*<sup>2</sup> of the graph is  $c(G) = 1$ .

The game proceeds in turns numbered by  $t \in \mathbb{N}_0$  and, as already mentioned, involves three tokens:  $C_1, C_2$  and  $R$ . These will also be referred to as the first, second and third token, respectively, and their locations *at the end* of the  $t$ -th turn are indicated by  $X_t^1, X_t^2, X_t^3$ . The *starting position* at the 0-th turn is *given*: for  $i \in \{1, 2, 3\}$ ,  $X_0^i = x_0^i \in V$ . In subsequent turns, the positions are changed according to the rules of the particular variant (sequential or concurrent) and subject to the constraint that movement always follows the graph edges:  $X_{t+1}^i \in N[X_t^i]$  (the *closed* neighborhood of  $X_t^i$ ). As will be seen in the sequel, in the general case token moves are governed by *probabilistic strategies*; hence  $X_t^1, X_t^2, X_t^3$  are random variables.

### 2.1 Sequential SCPR

In the *sequential* version of SCPR players take turns in moving their tokens. More specifically, on odd-numbered turns  $C_1$  is moved by the first cop player; on even-numbered turns first  $C_2$  is moved by the second cop player and then  $R$  is moved by Chance. Consequently, for  $t = 2l + 1$  we have  $X_t^2 = X_{t-1}^2$  and  $X_t^3 = X_{t-1}^3$ ; for  $t = 2l$  we have  $X_t^1 = X_{t-1}^1$ . An additional sequence of variables  $U_0, U_1, U_2, U_3, \dots$  indicates the player to move in the next turn; i.e.,  $U_0 = U_2 = \dots = 1, U_1 = U_3 = \dots = 2$ . We also define the vector  $S_t = (X_t^1, X_t^2, X_t^3, U_t)$ .

A *game position* or *state* is a vector  $s = (x^1, x^2, x^3, u)$  where  $x^1, x^2, x^3$  are the positions of the three tokens and  $u$  indicates which cop is about to play. For instance,  $s = (2, 3, 5, 1)$  denotes the situation in which  $C_1, C_2, R$  are located at vertices 2, 3, 5, respectively, and  $C_1$  will move in the next turn. We define the following sets of states

$$\forall i \in \{1, 2\} : \mathbf{S}_i = V \times V \times V \times \{i\}.$$

In other words,  $\mathbf{S}_1$  (resp.  $\mathbf{S}_2$ ) is the set of states “belonging” to the first (resp. second) player. A  $C_1$ -*capture state* is an  $s = (x^1, x^2, x^3, u)$  such that  $x^1 = x^3$ . A  $C_2$ -*capture state*<sup>3</sup> is an  $s = (x^1, x^2, x^3, u)$  such that  $x^2 = x^3$  and  $x^1 \neq x^3$ . We will also use a *terminal* state, denoted by  $\tau$ ; the behavior of the terminal state will be described in detail a little later. At any rate, the full *state space* of the sequential SCPR game is

$$\mathbf{S} = \mathbf{S}_1 \cup \mathbf{S}_2 \cup \{\tau\}.$$

The random variable  $A_t^i$  denotes the *move* (or *action*) of the  $i$ -th token at time  $t$ . When the game state is  $s$ , the set of moves available to the  $i$ -th token is denoted by  $\mathbf{A}_i(s)$ . For instance, when  $s = (x^1, x^2, x^3, 1)$  we have  $\mathbf{A}_1(s) = N[x^1]$  (the *closed* neighborhood of  $x^1$ ) and  $\mathbf{A}_2(s) = \{x^2\}$ . Similar things hold for states  $s = (x^1, x^2, x^3, 2)$ . For  $s = \tau$  we have  $\mathbf{A}_i(s) = \{\lambda\}$ , where  $\lambda$  is the *null move*. Legal moves result to “normal” state transitions; e.g., suppose the current state is  $s = (2, 3, 5, 1)$  and the next moves are  $a^1 = 3, a^2 = 3, a^3 = 5$ ; then, assuming  $3 \in N[2]$ , the next state is  $s' = (3, 3, 5, 2)$ . However, the terminal state  $\tau$  raises the following exceptions.

1. If the current state  $s$  is a  $C_i$ -capture state ( $i \in \{1, 2\}$ ), then the next state is  $s' = \tau$ , irrespective of the token moves. In other words, a capture state always transits to the terminal state.
2. If the current state  $s$  is the terminal (i.e.,  $s = \tau$ ), then the next state is  $s' = \tau$  irrespective of the token moves. In other words, the terminal always transits to itself.

A *play* or *infinite history* of the SCPR game is an infinite sequence  $s_0 s_1 s_2 \dots s_n \dots$  of game states. The set of all infinite histories is denoted by

$$H^\infty = \{s_0 s_1 s_2 \dots s_t \dots : s_t \in S \text{ for } t \in \mathbb{N}_0\}.$$

A *finite history* is a sequence  $s_0 s_1 s_2 \dots s_n$  of game states; the set of all histories of length  $n$  is denoted by

$$H_n = \{s_0 s_1 s_2 \dots s_{n-1} : s_t \in S \text{ for } t \in \{0, 1, \dots, n-1\}\};$$

<sup>2</sup>I.e., the minimum number of cops required to guarantee capture of the robber.

<sup>3</sup> $C_1$  is slightly favored, since an  $(x, x, x, u)$  state is considered a  $C_1$  capture; however, because of symmetry, reversing the definitions of  $C_i$ -captures would yield essentially the same results.

the set of all finite histories is  $H = \cup_{n=0}^{\infty} H_n$ .

We have already mentioned that each cop player moves his respective token. Rather than specifying each move separately, we assume (as is usual in Game Theory) that before the game starts, each cop player selects a *strategy* which controls all subsequent moves. Despite the fact that there is no robber player, we will assume that robber movement is also controlled by a “strategy”, which has been fixed before the game starts and *is known to the cop players*. Hence the  $i$ -th token ( $i \in \{1, 2, 3\}$ ) is controlled by the strategy (conditional probability function):

$$\sigma_i(a|s_0s_1\dots s_t) = \Pr(A_{t+1}^i = a | (X_0^1, X_0^2, X_0^3, U_0) = s_0, \dots, (X_t^1, X_t^2, X_t^3, U_t) = s_t).$$

The above definition is sufficiently general to describe every possible manner of move selection. We will only consider strategies which assign zero probability to illegal moves. The following classes of strategies are of particular interest.

1. A strategy  $\sigma_i$  is called *stationary Markovian* (or *positional*) iff  $\sigma_i(a|s_0s_1\dots s_t) = \sigma_i(a|s_t)$ ; i.e., the probability of the next move depends only on the current *state* of the game.
2. A strategy  $\sigma_i$  is called *oblivious* iff it is stationary Markovian and  $\sigma_i(a|(y^1, y^2, y^3, u)) = \sigma_i(a|y^i, u)$ ; i.e., the probability of the next move of the token depends only on (i) the current *location* of the token and (ii) the active player.
3. A strategy  $\sigma_i$  is called *deterministic* iff, for every  $s_0s_1\dots s_t \in H$ ,  $\sigma_i(x|s_0s_1\dots s_t) \in \{0, 1\}$ ; i.e., for every history, the  $i$ -th token moves to its next location deterministically.

To simplify presentation, we will often use the following notation for deterministic strategies. We define the *deterministic strategy* to be a function  $\bar{\sigma}_i : H \rightarrow V$ , defined as follows: for every finite history  $s_0s_1\dots s_t$ ,  $\bar{\sigma}_i(s_0s_1\dots s_t) = a$ , where  $a$  is the unique vertex such that  $\sigma_i(a|s_0s_1\dots s_t) = 1$ . If  $\sigma_i$  is stationary Markovian, we write  $\bar{\sigma}_i(s_t) = a$ .

Suppose the game is in state  $s$ . Now  $C_1$  plays  $a^1$ ,  $C_2$  plays  $a^2$  and  $R$ 's move  $a^3$  is selected according to the (fixed) strategy  $\sigma_3$ ; hence the game will move into some new state  $s'$  with a certain probability depending on  $a_1, a_2$  and  $\sigma_3$ . We denote this probability by  $\Pr(s'|s, a_1, a_2)^4$ .

Payoff is defined as follows. In each turn of the game,  $C_1$  receives an *immediate payoff* equal to

$$q(s) = \begin{cases} 1 & \text{iff } s \text{ is a } C_1\text{-capture state} \\ 0 & \text{otherwise} \end{cases}; \quad (1)$$

$C_2$  receives  $-q(s)$ . Hence, a play  $s_0s_1\dots$  results in (total) *payoff*

$$Q(s_0s_1\dots) = \sum_{t=0}^{\infty} q(s_t) \quad (2)$$

for  $C_1$  and  $-Q(s_0s_1\dots)$  for  $C_2$ . Note that both players have an incentive to capture  $R$ .

1. If  $C_1$  captures the robber, he receives a total payoff of one (comprising of immediate payoff of one for the capture turn and zero for all other turns); otherwise his total payoff is zero.
2.  $C_2$  *never* receives positive payoff (even if he captures the robber). However, we have assumed  $c(G) = 1$  and this implies that a single cop can always catch the robber. Hence, if  $C_2$  does not capture  $R$ ,  $C_1$  will and thus  $C_2$  will receive a negative payoff; this provides the incentive for  $C_2$  to capture  $R$ .

Sequential SCPR is a *stochastic zero sum game* [13]. Each player will try to maximize his *expected* payoff. Suppose the game starts at position  $s_0$ ,  $C_i$  moves according to strategy  $\sigma_i$  (for  $i \in \{1, 2\}$ ) and  $R$  moves according to a fixed and known strategy  $\sigma_3$ . Every triple  $(\sigma_1, \sigma_2, \sigma_3)$  induces a probability measure on  $H^\infty$ , the set of all infinite game histories. Hence the *expected* payoff to  $C_1$  is

$$J(\sigma_1, \sigma_2|s_0) = \mathbb{E} \left( \sum_{t=0}^{\infty} q(s_t) \mid (X_0^1, X_0^2, X_0^3, U_0) = s_0 \right) \quad (3)$$

---

<sup>4</sup>Note that in all subsequent notation, the dependence on the fixed and known  $\sigma_3$  is suppressed. Also, when a cop reaches the vertex occupied by the robber we have a capture with probability one, irrespective of the robber's move.

and is well defined;  $-J(\sigma_1, \sigma_2|s_0)$  is the expected payoff to  $C_2$ . It is easily seen that

$$J(\sigma_1, \sigma_2|s_0) = \Pr(\text{"}C_1 \text{ wins"} | \text{"the game starts at } s_0 \text{ and, for } i \in \{1, 2\}, C_i \text{ uses } \sigma_i").$$

We always have

$$\sup_{\sigma_1} \inf_{\sigma_2} J(\sigma_1, \sigma_2|s_0) \leq \inf_{\sigma_2} \sup_{\sigma_1} J(\sigma_1, \sigma_2|s_0); \quad (4)$$

if the two sides of (4) are equal, we define the *value* of the game (when started at  $s_0$ ) to be

$$v(s_0) = \sup_{\sigma_1} \inf_{\sigma_2} J(\sigma_1, \sigma_2|s_0) = \inf_{\sigma_2} \sup_{\sigma_1} J(\sigma_1, \sigma_2|s_0); \quad (5)$$

$\mathbf{v}$  will denote the vector of values for all starting states, i.e.,  $\mathbf{v} = (v(s))_{s \in \mathbf{S}}$ . Given some  $\varepsilon \geq 0$ , we say that:

1. a strategy  $\sigma_1^\#$  is  $\varepsilon$ -*optimal* (for  $C_1$ ) iff  $\forall s_0 : v(s_0) - \inf_{\sigma_2} J(\sigma_1^\#, \sigma_2|s_0) \leq \varepsilon$ ;
2. a strategy  $\sigma_2^\#$  is  $\varepsilon$ -*optimal* (for  $C_2$ ) iff  $\forall s_0 : v(s_0) - \sup_{\sigma_1} J(\sigma_1, \sigma_2^\#|s_0) \geq -\varepsilon$ .

A 0-optimal strategy is also called simply *optimal*.

Finally, let  $\Gamma$  be a *matrix game*, i.e., a (one-turn) two-player, zero-sum game with finite action set  $\mathbf{A}_i$  for the  $i$ -th player and the payoff to the first player being  $\Gamma(a^1, a^2)$  when  $i$ -th player plays  $a^i \in \mathbf{A}_i$  (with  $i \in \{1, 2\}$ ). As is well known [22], such a game always has a *value*, which we will denote by  $\mathbf{Val}[\Gamma(a^1, a^2)]$ .

## 2.2 Concurrent SCPR

Most of the CR literature studies sequential versions of the CR game. However, we have recently introduced a *concurrent* version of the classic CR [15]. Now we extend concurrency to the SCPR game.

The concurrent SCPR game differs from the sequential game in a basic respect: in every turn the  $C_1, C_2, R$  tokens are moved *simultaneously* (hence, when making his move, each player does not know the other player's move; note that both of them know  $R$ 's next move, since  $\sigma_3$  is known in advance). Once again we will assume, unless otherwise indicated, that  $\hat{c}(G) = 1$  (note that a graph  $G$  has *concurrent cop number*  $\hat{c}(G) = k$  iff it has sequential, i.e., "classic", cop number  $c(G) = k$  [15]).

In addition, in concurrent SCPR we can have "*en-passant capture*", in which a cop and the robber start at opposite ends of the same edge and move in opposite directions; in this case the robber is "swept" by the cop and moved into the cop's destination.

With concurrent movement, game states are vectors  $(x^1, x^2, x^3)$  where  $x^i \in V$  indicates (as previously) the position of the  $i$ -th token; the  $u$  variable is no longer necessary, since all tokens are moved in every turn. Capture states now have the form  $(x^1, x^2, x^3)$  with either  $x^1 = x^3$  or  $x^2 = x^3$  (or both) and the definition and behavior of the terminal state  $\tau$  are the same as previously. For the state space, we define

$$\hat{\mathbf{S}}_a = V \times V \times V, \quad \hat{\mathbf{S}} = \hat{\mathbf{S}}_a \cup \{\tau\}$$

and  $\hat{\mathbf{S}}$  is the full *state space* of the of concurrent SCPR game.

Regarding  $\mathbf{A}_i(s)$  (the actions available to the  $i$ -th player when the game state is  $s$ ) we always have  $\mathbf{A}_i((x^1, x^2, x^3)) \in N[x^i]$ . The definitions of (finite and infinite) histories and strategies are the same as in the sequential case, except that we now use the state space  $\hat{\mathbf{S}}$ . The meaning of the sets  $\hat{H}_n, \hat{H}, \hat{H}^\infty$  is analogous to that of  $H_n, H, H^\infty$ . The strategies  $\sigma_i$  ( $i \in \{1, 2, 3\}$ ) are defined in the same manner as in the sequential case (again, for deterministic moves we introduce the deterministic strategy functions  $\bar{\sigma}_i$ ).

Payoff of the concurrent SCPR game is defined in exactly the same manner as in the sequential case. Again, concurrent SCPR is a stochastic zero sum game and each player will try to maximize his expected payoff.

## 3 Sequential SCPR

In this section we establish that sequential SCPR has a *value* which can be computed by *value iteration*.

**Theorem 3.1** *Given some graph  $G = (V, E)$ . For every  $s \in \mathbf{S}_1 \cup \mathbf{S}_2$ , the sequential SCPR game starting at  $s$  has a value  $v(s)$ . The vector of values  $\mathbf{v} = (v(s))_{s \in \mathbf{S}}$  is the smallest (componentwise) solution of the following optimality equations:*

$$v(\tau) = 0; \quad (6)$$

$$\forall s = (x^1, x^2, x^3, 1) \in \mathbf{S}_1 : \quad v(s) = \max_{a^1} \left[ q(s) + \sum_{s' \in \mathbf{S}} \Pr(s'|s, a^1, x^2) v(s') \right], \quad (7)$$

$$\forall s = (x^1, x^2, x^3, 2) \in \mathbf{S}_2 : \quad v(s) = \min_{a^2} \left[ q(s) + \sum_{s' \in \mathbf{S}} \Pr(s'|s, x^1, a^2) v(s') \right]. \quad (8)$$

Furthermore  $C_2$  has a deterministic stationary Markovian optimal strategy and, for every  $\varepsilon > 0$ ,  $C_1$  has a deterministic stationary Markovian  $\varepsilon$ -optimal strategy.

**Proof.** It is easily checked that, for every graph  $G$  and every starting position  $s$ , the sequential SCPR game is a positive zero sum stochastic game. Hence (by [13, Theorem 4.4.1]) it possesses a value which (by [13, Theorem 4.4.3]) is the smallest componentwise solution to the following system of *optimality equations*:

$$v(\tau) = 0; \quad \forall s \in \mathbf{S}_1 \cup \mathbf{S}_2 : v(s) = \mathbf{Val} \left[ q(s) + \sum_{s' \in \mathbf{S}} \Pr(s'|s, a^1, a^2) v(s') \right]. \quad (9)$$

However, in each turn of the sequential SCPR game, one of the players has a single available action. For instance, when the state is  $s = (x^1, x^2, x^3, 1)$ ,  $C_2$ 's action set can only be  $a^2 = x^2$ . Hence in (9) we are taking the value of an one-shot game with the game matrix consisting of a single column. It follows that

$$\forall s = (x^1, x^2, x^3, 1) : \mathbf{Val} \left[ q(s) + \sum_{s' \in \mathbf{S}} \Pr(s'|s, a^1, a^2) v(s') \right] = \max_{a^1} \left[ q(s) + \sum_{s' \in \mathbf{S}} \Pr(s'|s, a^1, x^2) v(s') \right]$$

which proves (7); (8) can be proved similarly.

The existence of stationary Markovian optimal strategy for  $C_2$  follows from [13, Corollary 4.4.2]. It is a deterministic strategy because for each state  $s \in \mathbf{S}_2$  the corresponding optimal  $C_2$  move is the one minimizing (8). Similarly, the existence of a stationary Markovian  $\varepsilon$ -optimal strategy for  $C_1$  follows from [13, Problem 4.16]; the strategy is deterministic, because for each state  $s \in \mathbf{S}_1$  the corresponding optimal  $C_1$  move is the one maximizing (7). ■

For the computation of the solution to (7)-(8) we have the following.

**Theorem 3.2** *Given some graph  $G = (V, E)$ . Define  $\mathbf{v}^{(0)}$  by*

$$v^{(0)}(\tau) = 0; \quad \forall s \in \mathbf{S}_1 \cup \mathbf{S}_2 : v^{(0)}(s) = q(s)$$

*and  $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots$  by the following recursion*

$$v^{(i)}(\tau) = 0 \quad (10)$$

$$\forall s = (x^1, x^2, x^3, 1) \in \mathbf{S}_1 : \quad v^{(i)}(s) = \max_{a^1} \left[ q(s) + \sum_{s' \in \mathbf{S}} \Pr(s'|s, a^1, x^2) v^{(i-1)}(s') \right], \quad (11)$$

$$\forall s = (x^1, x^2, x^3, 2) \in \mathbf{S}_2 : \quad v^{(i)}(s) = \min_{a^2} \left[ q(s) + \sum_{s' \in \mathbf{S}} \Pr(s'|s, x^1, a^2) v^{(i-1)}(s') \right]. \quad (12)$$

*Then, for every  $s \in \mathbf{S}_1 \cup \mathbf{S}_2$ ,  $\lim_{i \rightarrow \infty} v^{(i)}(s)$  exists and equals  $v(s)$ , the value of the sequential SCPR game played on  $G$ , starting from  $s$ .*

**Proof.** Obviously, for all  $s \in \mathbf{S}$ ,  $v(s) \in [0, 1]$ . Hence  $\mathbf{v}$  is a (componentwise) finite vector. Then from [13, Theorem 4.4.4] we know that, defining  $\mathbf{v}^{(0)}$  by

$$v^{(0)}(\tau) = 0; \quad \forall s \in \mathbf{S}_1 \cup \mathbf{S}_2 : v^{(0)}(s) = q(s)$$

and  $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots$  by the recursion

$$v^{(i)}(\tau) = 0; \quad \forall s \in \mathbf{S}_1 \cup \mathbf{S}_2 : \quad v^{(i)}(s) = \mathbf{Val} \left[ q(s) + \sum_{s' \in S} \Pr(s'|s, a_1, a_2) v^{(i-1)}(s') \right], \quad (13)$$

we get  $\lim_{i \rightarrow \infty} \mathbf{v}^{(i)} = \mathbf{v}$  (the value vector of Theorem 3.1). The equivalence of (13) to (10)-(12) is established by the argument used in the proof of Theorem 3.1. ■

**Remark 3.3** The significance of Theorem 3.1 is the following. Since SCPR is a positive zero sum stochastic game, it will certainly have a value, which satisfies the optimality equations (9); each equation of the system (9) involves the value of a one-turn game. However, the optimality equations can be expressed in the simpler form (7)-(8) which show that the values of the one-step games can be computed by simple max and min operations.

**Remark 3.4** Similar remarks can be made about Theorem 3.2, where the iteration (10)-(12) is computationally simpler (involves only max and min operations) than (13). Note the similarity of (10)-(12) to the algorithm of [14] for determining the winner of a classic CR game. The similarity becomes stronger in the case of deterministic  $\sigma_3$ . In this case,  $\Pr(s'|s, a^1, x^2)$  equals 1 for a single  $s' = \mathbf{T}(s, a^1, x^2)$  and  $\Pr(s'|s, x^1, a^2)$  equals 1 for a single  $s' = \mathbf{T}(s, x^1, a^2)$ ; where  $\mathbf{T}(s, a^1, a^2)$  is the *transition function* which yields the next state when, from  $s$ ,  $C_1$  plays  $a^1$  and  $C_2$  plays  $a^2$ ; there is also a suppressed dependence on the move of  $R$ , which is  $\bar{\sigma}_3(s)$ . Using this notation, (10)-(12) simplify to

$$\forall s = (x^1, x^2, x^3, 1) \in \mathbf{S}_1 : v^{(i)}(s) = \max_{a^1} \left[ q(s) + v^{(i-1)}(\mathbf{T}(s, a^1, x^2)) \right], \quad (14)$$

$$\forall s = (x^1, x^2, x^3, 2) \in \mathbf{S}_2 : v^{(i)}(s) = \min_{a^2} \left[ q(s) + v^{(i-1)}(\mathbf{T}(s, x^1, a^2)) \right]; \quad (15)$$

these parallel closely the algorithm of [14, p.2494].

**Remark 3.5** Finally, note that Theorems 3.1 and 3.2 hold even when  $c(G) > 1$ ; the reason for which we have previously required  $c(G) = 1$  has to do with the appropriateness of the payoff function introduced in Section 2. In particular, when  $c(G) > 1$  our argument about  $C_2$ 's incentive to capture  $R$  does not hold necessarily (i.e., depending on  $\sigma_3$ ,  $C_2$  may ensure payoff of 0 without ever capturing  $R$ ); Theorems 3.1 and 3.2 still hold true.

## 4 Concurrent SCPR

In this section we establish that concurrent SCPR has a value which can be computed by value iteration. We first consider the case in which  $R$  is controlled by a general probability function  $\sigma_3$  ("random robber") and then examine in greater detail the case in which  $\sigma_3$  is oblivious deterministic ("oblivious deterministic robber").

### 4.1 Random Robber

The two main results on concurrent SCPR are immediate consequences of the more general results of [13].

**Theorem 4.1** *Given some graph  $G = (V, E)$ . For every  $s = (x^1, x^2, x^3) \in \widehat{\mathbf{S}}_a$ , the concurrent SCPR game starting at  $s$  has a value  $v(s)$ . The vector of values  $\mathbf{v} = (v(s))_{s \in \mathbf{S}}$  is the smallest (componentwise) solution of the following optimality equations*

$$v(\tau) = 0; \quad \forall s \in \widehat{\mathbf{S}}_a : v(s) = \mathbf{Val} \left[ q(s) + \sum_{s' \in S} \Pr(s'|s, a_1, a_2) v(s') \right]. \quad (16)$$

Furthermore,  $C_2$  has a stationary Markovian optimal strategy and, for every  $\varepsilon > 0$ ,  $C_1$  has a stationary Markovian  $\varepsilon$ -optimal strategy.

**Proof.** For every graph  $G$  (and every starting position  $s$ ) SCPR is a positive stochastic game. Hence (by [13, Theorem 4.4.1]) it possesses a value which (by [13, Theorem 4.4.3]) satisfies the optimality equation (16). Furthermore  $C_2$  has a stationary Markovian optimal strategy by [13, Corollary 4.4.2] and, for every  $\varepsilon > 0$ ,  $C_1$  has a stationary Markovian  $\varepsilon$ -optimal strategy by [13, Problem 4.16]. ■

**Theorem 4.2** Given some graph  $G = (V, E)$ , let  $s = (x^1, x^2, x^3) \in \widehat{\mathbf{S}}_a$ . Define  $\mathbf{v}^{(0)}$  by

$$v^{(0)}(\tau) = 0; \quad \forall s \in \widehat{\mathbf{S}}_a : v^{(0)}(s) = q(s)$$

and  $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots$  by the following recursion

$$v^{(i)}(\tau) = 0; \quad \forall s \in \widehat{\mathbf{S}}_a : v^{(i)}(s) = \mathbf{Val} \left[ q(s) + \sum_{s' \in \mathbf{S}} \Pr(s'|s, a_1, a_2) v^{(i-1)}(s') \right]. \quad (17)$$

Then, for every  $s \in \widehat{\mathbf{S}}_a$ ,  $\lim_{i \rightarrow \infty} v^{(i)}(s)$  exists and equals  $v(s)$ , the value of the concurrent SCPR game played on  $G$ , starting from  $s$ .

**Proof.** This follows immediately from [13, Theorem 4.4.4]. ■

## 4.2 Oblivious Deterministic Robber

Theorems 3.1 and 3.2 are “simpler” than Theorems 4.1 and 4.2, in the sense that the former do not involve the computation of matrix game values. We will now show that, when  $\sigma_3$  is oblivious deterministic, we can obtain a similar simplification of Theorems 4.1 and 4.2. Before presenting these results in rigorous form, let us describe them informally.

1. Suppose first that a game is played between a single cop and an oblivious deterministic robber. We will prove that there exists a stationary Markovian deterministic cop strategy  $\bar{\sigma}^*$  by which the cop can capture the robber in minimum time.
2. Next consider two cops and an oblivious deterministic robber. We will prove that the extension of  $\bar{\sigma}^*$  to SCPR is *optimal* for both cops. More specifically, neither cop loses anything by using it; and one of the two will capture the robber with probability one.

Let us now formalize the above ideas. We pick any graph  $G = (V, E)$  and any oblivious deterministic robber strategy  $\bar{\sigma}_3$  and keep these fixed for the remainder of the discussion. Further, let  $\mathcal{S}$  denote the set of all functions  $\bar{\sigma} : V \times V \rightarrow V$  with the restriction that  $\forall (x^1, x^3) \in V \times V : \bar{\sigma}(x^1, x^3) \in N[x^1]$ . In other words,  $\mathcal{S}$  is the set of legal stationary Markovian deterministic cop strategies for the “classic” CR game of one cop and one robber.

Now pick some  $\bar{\sigma} \in \mathcal{S}$  and play the game with starting positions  $X_0^1 = x_0^1 \in V$  (for the cop) and  $X_0^3 = x_0^3 \in V$  (for the robber). The following sequence (dependent on  $\bar{\sigma}, x_0^1, x_0^3$ ) of cop and robber positions will be produced:

$$X_0^1 = x_0^1, X_0^3 = x_0^3, X_1^1 = \bar{\sigma}(x_0^1, x_0^3), X_1^3 = \bar{\sigma}_3(x_0^3), \dots;$$

let  $T_{\bar{\sigma}}(x_0^1, x_0^3)$  be the *capture time*, i.e., the smallest  $t$  such that  $X_t^1 = X_t^3$ , for the sequence produced by  $\bar{\sigma}, x_0^1, x_0^3$  (and  $\bar{\sigma}_3$ ). Also define

$$\overline{V \times V} = \{(x^1, x^3) : x^1 \in V, x^3 \in V, x^1 \neq x^3\}.$$

Then we have the following.

**Lemma 4.3** Given a graph  $G = (V, E)$  and an oblivious deterministic robber strategy  $\bar{\sigma}_3$ . Let

$$\forall x^1 \in V : T^{(0)}(x^1, x^1) = 0, \quad \forall (x^1, x^3) \in \overline{V \times V} : T^{(0)}(x^1, x^3) = \infty.$$

Now perform the following iteration for  $i = 1, 2, \dots$ :

$$\forall x^1 \in V : T^{(i)}(x^1, x^1) = 0; \quad \forall (x^1, x^3) \in \overline{V \times V} : T^{(i)}(x^1, x^3) = \min_{x' \in N[x^1]} \left[ 1 + T^{(i-1)}(x', \sigma_3(x^3)) \right], \quad (18)$$

$$\forall x^1 \in V : T^{(i)}(x^1, x^1) = 0; \quad \forall (x^1, x^3) \in \overline{V \times V} : \bar{\sigma}^{(i)}(x^1, x^3) = \arg \min_{x' \in N[x^1]} \left[ 1 + T^{(i-1)}(x', \sigma_3(x^3)) \right]. \quad (19)$$

Then the limits

$$\lim_{i \rightarrow \infty} \bar{\sigma}^{(i)}(x^1, x^3), \quad \lim_{i \rightarrow \infty} T^{(i)}(x^1, x^3)$$

exist for all  $(x^1, x^3) \in V \times V$ . Furthermore, letting  $\bar{\sigma}^*(x^1, x^3) = \lim_{i \rightarrow \infty} \bar{\sigma}^{(i)}(x^1, x^3)$  and  $T^*(x^1, x^3) = \min_{\bar{\sigma} \in \mathcal{S}} T_{\bar{\sigma}}(x^1, x^3)$ , we have

$$\forall (x^1, x^3) \in V \times V : \lim_{i \rightarrow \infty} T^{(i)}(x^1, x^3) = T_{\bar{\sigma}^*}(x^1, x^3) = T^*(x^1, x^3). \quad (20)$$

**Proof.** The proof is based on a standard dynamic programming argument. First note that, for every  $(x^1, x^3) \in V \times V$ ,  $T^*(x^1, x^3) < |V|$ . This is true because  $C_1$  can reach any vertex of  $V$  in at most  $|V| - 1$  moves; so  $C_1$  can simply go to  $X_{|V|}^3$  (the *known* location of  $R$  at time  $t = |V|$ ) and wait for the robber there.

Next we prove by induction that

$$T^*(x^1, x^3) = n \Rightarrow \left( \forall i \geq n : T^*(x^1, x^3) = T^{(i)}(x^1, x^3) \right). \quad (21)$$

For  $n = 0$ ,  $T^*(x^1, x^3) = 0$  implies  $x^1 = x^3$  and, from the algorithm,  $T^*(x^1, x^1) = 0 = T^{(i)}(x^1, x^1)$  for all  $i \in \mathbb{N}_0$ . Now suppose that (21) holds for  $n = 1, 2, \dots, k$  and consider the case  $n = k + 1$ , in which  $T^*(x^1, x^3) = k + 1$  is the smallest number of steps in which  $C_1$  can reach  $R$ . This also means that (i) there exists some  $x' \in N[x^1]$  from which  $C_1$  can reach  $R$  (who now starts at  $\bar{\sigma}_3(x^3)$ ) in  $k$  steps and (ii) there does not exist any  $x'' \in N[x^1]$  from which  $C_1$  can reach  $R$  in  $m < k$  steps (because then  $C_1$  starting at  $x^1$  could reach  $R$  in  $m + 1 < k + 1$  steps). In other words

$$T^*(x^1, x^3) = k + 1 \Rightarrow T^*(x^1, x^3) = \min_{x' \in N[x^1]} \left[ 1 + T^{(k)}(x', \bar{\sigma}_3(x^3)) \right] = T^{(k+1)}(x^1, x^3).$$

It is also easy to check that:

$$\forall m \in \mathbb{N}_0 : T^{(m)}(x^1, x^3) = m \Rightarrow \left( \forall i > m : T^{(i)}(x^1, x^3) = m \right).$$

Hence the induction has been completed.

Given (21), we see immediately that

$$\forall (x^1, x^3) \in V \times V, i \geq |V| : T^{(i)}(x^1, x^3) = T^*(x^1, x^3)$$

which implies that both  $\lim_{i \rightarrow \infty} T^{(i)}(x^1, x^3) = T^*(x^1, x^3)$  and  $\lim_{i \rightarrow \infty} \sigma^{(i)}(x^1, x^3)$  exist. Taking the limit (as  $i$  tends to  $\infty$ ) in (18)-(19) we get the *optimality equations*

$$\begin{aligned} T^*(x^1, x^3) &= \min_{x' \in N[x^1]} [1 + T^*(x', \bar{\sigma}_3(x^3))] \\ \bar{\sigma}^*(x^1, x^3) &= \arg \min_{x' \in N[x^1]} [1 + T^*(x', \bar{\sigma}_3(x^3))] \end{aligned}$$

hence, it is clear from the iteration (18)-(19) that  $T_{\bar{\sigma}^*}(x^1, x^3) = T^*(x^1, x^3)$ , for all  $(x^1, x^3) \in V \times V$ . ■

Now let us use  $\bar{\sigma}^*$  of Lemma 4.3 to define strategies  $\bar{\sigma}_i^*$  for  $C_i$  ( $i \in \{1, 2\}$ ) as follows:

$$\begin{aligned} \forall (x^1, x^2, x^3) \in \mathbf{S}_a : \bar{\sigma}_1^*(x^1, x^2, x^3) &= \bar{\sigma}^*(x^1, x^3), \\ \forall (x^1, x^2, x^3) \in \mathbf{S}_a : \bar{\sigma}_2^*(x^1, x^2, x^3) &= \bar{\sigma}^*(x^2, x^3). \end{aligned}$$

Then the following holds.

**Theorem 4.4** *Given some graph  $G = (V, E)$ , suppose SCPR is played on  $G$  and the robber is controlled by an oblivious deterministic strategy  $\bar{\sigma}_3$ . Then  $\bar{\sigma}_i^*$  is an optimal strategy for  $C_i$  ( $i \in \{1, 2\}$ ), for every starting position  $s = (x^1, x^2, x^3) \in \mathbf{S}_a$ . Furthermore*

$$\forall s = (x^1, x^2, x^3) \in \mathbf{S}_a : \begin{aligned} T_{\bar{\sigma}_1^*}(x^1, x^3) &\leq T_{\bar{\sigma}_2^*}(x^2, x^3) &\Rightarrow v(s) &= 1 \\ T_{\bar{\sigma}_1^*}(x^1, x^3) &> T_{\bar{\sigma}_2^*}(x^2, x^3) &\Rightarrow v(s) &= 0 \end{aligned}.$$



**Proof.** The key fact is this: when  $\bar{\sigma}_3$  is oblivious deterministic, the players  $C_1$  and  $C_2$  interact only at the last phase of the game, when  $R$  is captured. In effect each cop plays a “decoupled” classic CR game, in which  $\bar{\sigma}^*$  of Lemma 4.3 guarantees capture in minimum time. Of course in the full SCPR game there is always the possibility that the other cop can capture  $R$  at an earlier time. Hence the best  $C_i$  can do is to attempt to capture  $R$  at the earliest possible time and an optimal strategy to this end is  $\bar{\sigma}_i^*$ ; he has no incentive to deviate from  $\bar{\sigma}_i^*$  (by using another deterministic or probabilistic strategy) because this can never reduce his projected capture time. Hence  $\bar{\sigma}_i^*$  is optimal for  $C_1$ . Since  $\bar{\sigma}_1^*$ ,  $\bar{\sigma}_2^*$  and  $\bar{\sigma}_3$  are deterministic, the outcome of the game is also deterministic. In particular, when  $T_{\bar{\sigma}_1^*}(x^1, x^3) \leq T_{\bar{\sigma}_2^*}(x^2, x^3)$ , with probability 1  $C_1$  reaches  $R$  before or at the same time as  $C_2$ ; hence  $v(s) = 1$ ; when  $T_{\bar{\sigma}_1^*}(x^1, x^3) > T_{\bar{\sigma}_2^*}(x^2, x^3)$ ,  $C_2$  reaches  $R$  before  $C_1$  with probability 1; hence  $v(s) = 0$ . ■

The next theorem gives an additional characterization of the value  $v(s)$ . In the statement of the theorem we will use the following notation: suppose the game is in the state  $s$ ,  $C_1$  plays  $a^1$ ,  $C_2$  plays  $a^2$  and  $R$  plays the (predetermined) move  $\bar{\sigma}_3(s)$ ; then we denote the next game state by  $\hat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))$ . We have the following.

**Theorem 4.5** *Given some graph  $G = (V, E)$ , suppose SCPR is played on  $G$  and the robber is controlled by an oblivious deterministic strategy  $\bar{\sigma}_3$ . Then,  $\forall s \in \mathbf{S}_a$ , we have*

$$v(s) = \max_{a^1} \min_{a^2} \left[ q(s) + v\left(\hat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) \right] = \min_{a^2} \max_{a^1} \left[ q(s) + v\left(\hat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) \right]. \quad (22)$$

**Proof.** Since  $\bar{\sigma}_3$  is deterministic,  $\Pr\left(\hat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s))) \mid s, a^1, a^2\right) = 1$ . Hence, by [13, Theorem 4.4.3]:

$$v(s) = \mathbf{Val} \left[ q(s) + \sum_{s' \in S} \Pr(s' \mid s, a_1, a_2) v(s') \right] = \mathbf{Val} \left[ q(s) + v\left(\hat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) \right].$$

Since  $\bar{\sigma}_1^*$  and  $\bar{\sigma}_2^*$  are also deterministic, at every turn of the game they produce an action with probability one. Hence there exist actions  $\bar{a}^1 = \bar{\sigma}_1^*(s)$ ,  $\bar{a}^2 = \bar{\sigma}_2^*(s)$  such that

$$v(s) = q(s) + v\left(\hat{\mathbf{T}}(s, (\bar{a}^1, \bar{a}^2, \bar{\sigma}_3(s)))\right).$$

Since, from Theorem 4.4,  $v(s) \in \{0, 1\}$ , we consider two cases.

1. Suppose  $v(s) = 1$ . This means, that starting at  $s$ ,  $C_1$  will certainly capture  $R$ .

(a) If  $s$  is a  $C_1$ -capture state, then  $q(s) = 1$  and, for any actions  $\bar{a}^1, \bar{a}^2$ ,  $\hat{\mathbf{T}}(s, (\bar{a}^1, \bar{a}^2, \bar{\sigma}_3(s))) = \tau$ , in which case

$$v\left(\hat{\mathbf{T}}(s, (\bar{a}^1, \bar{a}^2, \bar{\sigma}_3(s)))\right) = v(\tau) = 0.$$

$$\text{Hence } v(s) = \max_{a^1} \min_{a^2} \left[ q(s) + v\left(\hat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) \right] = 1.$$

(b) If  $s$  is not a  $C_1$ -capture state, then  $q(s) = 0$  and  $v\left(\hat{\mathbf{T}}(s, (\bar{a}^1, \bar{a}^2, \bar{\sigma}_3(s)))\right) = 1$ . Suppose there existed some  $\hat{a}^2$  such that  $v\left(\hat{\mathbf{T}}(s, (\bar{a}^1, \hat{a}^2, \bar{\sigma}_3(s)))\right) = 0$ . This would mean that, starting at  $\hat{\mathbf{T}}(s, (\bar{a}^1, \hat{a}^2, \bar{\sigma}_3(s)))$ ,  $C_2$  would certainly capture  $R$  before  $C_1$  and, since  $\bar{a}^1$  is the optimal (fastest capturing) move for  $C_1$ , we would also have

$$\forall a^1 \in \mathbf{A}_1(s) : q(s) + v\left(\hat{\mathbf{T}}(s, (a^1, \hat{a}^2, \bar{\sigma}_3(s)))\right) = 0.$$

But then  $v(s) = \mathbf{Val} \left[ q(s) + v\left(\hat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) \right] = 0$ , contrary to the assumption. So we must instead have

$$\forall a^2 \in \mathbf{A}_2(s) : q(s) + v\left(\hat{\mathbf{T}}(s, (\bar{a}^1, a^2, \bar{\sigma}_3(s)))\right) = 1$$

which implies  $v(s) = \max_{a^1} \min_{a^2} \left[ q(s) + v\left(\hat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) \right] = 1$ .

2. Now suppose  $v(s) = 0$ . Then  $s$  is not a  $C_1$ -capture state, i.e.,  $q(s) = 0$ . Now, we will show that

$$\forall a^1 \in \mathbf{A}_1(s) : \exists a^2 \in \mathbf{A}_2(s) : v\left(\widehat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) = 0. \quad (23)$$

If this is not the case, then we must have

$$\exists \tilde{a}^1 \in \mathbf{A}_1(s) : \forall a^2 \in \mathbf{A}_2(s) : v\left(\widehat{\mathbf{T}}(s, (\tilde{a}^1, a^2, \bar{\sigma}_3(s)))\right) = 1.$$

Then  $C_1$  will certainly capture  $R$  (before  $C_2$ ) starting from the game position  $\mathbf{T}(s, (\tilde{a}^1, a^2, \bar{\sigma}_3(s)))$  and this will be true for any  $a^2 \in \mathbf{A}_2(s)$ . But this means that  $C_1$ , starting from game position  $s$  and playing  $\tilde{a}^1$ , will certainly capture  $R$  before  $C_2$ ; which in turn means  $v(s) = 1$ , contrary to the hypothesis. Hence (23) holds and this implies

$$\begin{aligned} \forall a^1 \in \mathbf{A}_1(s) : \min_{a^2} v\left(\widehat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) &= 0 \\ \Rightarrow \max_{a^1} \min_{a^2} v\left(\widehat{\mathbf{T}}(s, (a^1, a^2, \bar{\sigma}_3(s)))\right) &= 0. \end{aligned}$$

Hence we have proved the first part of (22). The proof of the second part is similar and omitted. ■

**Remark 4.6** It must be emphasized that Theorem 4.4 and Theorem 4.5 do *not* hold for deterministic *non*-oblivious strategies  $\bar{\sigma}_3$ . This can be seen by the following counterexample. Suppose that concurrent SCPR is played on the graph of Figure 1, starting from the state  $(2, 6, 1)$ .

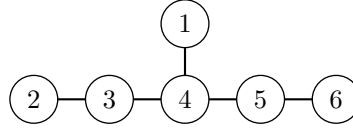


Figure 1: An example where deterministic robber strategy results in randomized optimal cop strategies.

Furthermore, the robber is controlled by the  $\bar{\sigma}_3$  which is (partially) described in the following table.

$(x_t^1, x_t^2, x_t^3)$	$x_{t+1}^3 = \bar{\sigma}_3(x_t^1, x_t^2, x_t^3)$
$(2, 6, 1)$	4
$(2, 6, 4)$	3
$(2, 5, 4)$	5
$(3, 6, 4)$	5
$(3, 5, 4)$	3

Table 1: A part of the robber strategy  $\sigma_3$

For every game state not listed above the robber stays in place, i.e.,  $x_{t+1}^3 = \bar{\sigma}_3(x_t^1, x_t^2, x_t^3) = x_t^3$ . Now consider what the first moves of  $C_1$  and  $C_2$  should be. They know that  $R$  will move into vertex 4;  $C_1$  can either stay at 2 or move into 3;  $C_2$  can either stay at 6 or move into 5. After the first move is completed, the possible game states are the following.

$s_0 = (2, 6, 1)$	$a_1^1 = 2$	$a_1^2 = 6$	$a_1^3 = \bar{\sigma}_3(2, 6, 1) = 4$	$s_1 = (2, 6, 4)$
$s_0 = (2, 6, 1)$	$a_1^1 = 2$	$a_1^2 = 5$	$a_1^3 = \bar{\sigma}_3(2, 6, 1) = 4$	$s_1 = (2, 5, 4)$
$s_0 = (2, 6, 1)$	$a_1^1 = 3$	$a_1^2 = 6$	$a_1^3 = \bar{\sigma}_3(2, 6, 1) = 4$	$s_1 = (3, 6, 4)$
$s_0 = (2, 6, 1)$	$a_1^1 = 3$	$a_1^2 = 5$	$a_1^3 = \bar{\sigma}_3(2, 6, 1) = 4$	$s_1 = (3, 5, 4)$

Table 2: Possible states at the end of the first turn.

It is easy to check (from the respective  $\bar{\sigma}_3$  values) that for  $s_1 = (2, 6, 4)$  and  $s_1 = (3, 5, 4)$  the capturing cop is  $C_1$ , while for  $s_1 = (2, 5, 4)$  and  $s_1 = (3, 6, 4)$  the capturing cop is  $C_2$ . Hence the game can be written out as the following (one-turn) matrix game

	$a^2 = 6$	$a^2 = 5$
$a^1 = 2$	1	0
$a^1 = 3$	0	1

Table 3: The one-turn matrix game equivalent to the original stochastic game.

It is easy to compute, using standard methods, that the optimal strategies for this game.  $C_1$  must use  $\Pr(a^1 = 2) = \Pr(a^1 = 3) = \frac{1}{2}$  and  $C_2$  must use  $\Pr(a^2 = 6) = \Pr(a^2 = 5) = \frac{1}{2}$ . This implies that the optimal strategies  $\bar{\sigma}_1^*$  and  $\bar{\sigma}_2^*$  are randomized, despite the fact that  $\bar{\sigma}_3$  is deterministic (but not oblivious). Many similar examples can be constructed. The important point is this: when  $\bar{\sigma}_3$  is not oblivious,  $C_1$  (resp.  $C_2$ ) moves can influence future  $R$  moves and (since moves are performed simultaneously) this influence cannot be predicted by  $C_2$  (resp.  $C_1$ ).

## 5 Related Work

In this section we present work which is related to both the SCPR and other variants of the CR game.

We have already mentioned that the interested reader can find useful references to the CR literature in the book [20] by Nowakowski and Bonato. The CR literature is mainly oriented to graph theoretic and combinatorial considerations. Indeed CR can be seen as a *combinatorial game*. On the topic of combinatorial games, the reader can consult the introductory text [1] as well as the classic book (in four volumes) [5] by Berlekamp and Conway. We also find interesting combinatorial generalizations of the CR game in the papers [7, 8] by A. Bonato and G. MacGillivray.

We believe that “classic” game theory offers a natural (but not often used in the “mainstream” CR literature) framework for the analysis of CR games. In particular, as already seen, we consider SCPR as a *stochastic game*. Stochastic games were introduced by Shapley [23]. A classic book on the subject is [13], which also contains a rich bibliography; see also [18].

A type of stochastic games which are especially related to CR games are *recursive games*, in which whenever a non-zero-payoff is received the play immediately moves to an absorbing state. Recursive games were introduced by Everett [12]. It is obvious that SCPR is a recursive game; while we have not used results from the recursive game theory in the current paper, we believe this may turn out to be a fruitful connection.

Let us now mention a construction which has been used in several “classic” CR papers [7, 8, 14]. Suppose that a “classic” CR game is played between one cop and one robber on the undirected graph  $G = (V, E)$ . We now construct the *game digraph*  $D = (S, A)$ , where the vertex set is  $S = V \times V \times \{1, 2\}$  and the arc set  $A$  encodes possible vertex-to-vertex transitions. Then a play of the CR game can be understood as a walk on  $D$ ; the cop wins if he can force the walk to pass through a vertex of the form  $(x, x, i)$ . Hence CR can be seen as a game in which the two players push a token along the arcs of the digraph. As pointed out in [7, 8] many CR variants and several other pursuit games on graphs (including the concurrent CR game) can be formulated in a similar manner.

It turns out that such “digraph games” have been studied by several researchers and the related literature is spread among many communities. The earliest such works of which we are aware is [4, 16]. Other early examples of this literature are the papers [3, 11, 24]. But probably the most widespread application of this point of view is in the literature of reachability games [6] and, more generally,  $\omega$ -regular games [17]. In a reachability game two players take turns moving a token along the arcs of a digraph; player 1 wants to place the token on one of the nodes of a subset of the digraph vertices while player 2 wants to avoid this event. In addition to “classic” sequential reachability games, many other variants have been studied, e.g., stochastic [9], concurrent [2],  $n$ -player [10] etc. The connection to CR games is obvious; it seems likely that the voluminous literature on reachability games contains results of interest to CR researchers.

## 6 Conclusion

We have introduced the game of selfish cops and passive robber (SCPR game) and established its basic properties, namely the existence of value and optimal strategies for both the sequential and concurrent variants; we have also provided algorithms for the computation of the aforementioned quantities. In the current paper we have examined *qualitative* variants of the game, i.e., these in which the goal of the cops is simply to capture the robber. In a forthcoming paper we will examine *quantitative* variants, in which the goal is to capture the robber *in the shortest possible time*.

Several additional issues merit further study and will be the subject of our future research. We have formulated SCPR as a zero-sum game; but reasonable formulations as a *non-zero-sum* game are also possible and we conjecture that these may lead to qualitatively different results. In addition, if we remove the assumption that the robber is passive and deal instead with the situation of two selfish robbers and a robber *actively trying to avoid capture*, we are left with a *three-player* game, which we intend to study in the future.

## References

- [1] M.H. Albert, R. Nowakowski and D. Wolfe (2007). *Lessons in Play: An Introduction to Combinatorial Game Theory*.
- [2] L. de Alfaro, T. A. Henzinger and O. Kupferman. “Concurrent reachability games”. *Theoretical Computer Science*, vol.386 (2007), pp. 188-217.
- [3] V. J. Baston and F. A. Bostock. “Infinite deterministic graphical games”. *SIAM journal on control and optimization* vol.31 (1993). pp.1623-1629.
- [4] A. Berarducci and B. Intrigila. “On the cop number of a graph”. *Advances in Applied Mathematics*, vol. 14 (1993), pp. 389-403.
- [5] E. Berlekamp, J. H. Conway and R. Guy (1982). *Winning Ways for your Mathematical Plays*.
- [6] D. Berwanger, Graph games with perfect information, *preprint*.
- [7] A. Bonato and G. MacGillivray. “A general framework for discrete - time pursuit games”, *preprint*.
- [8] A. Bonato and G. MacGillivray. “Characterizations and algorithms for generalized Cops and Robbers games”, accepted to *Contributions to Discrete Mathematics* (2016).
- [9] K. Chatterjee and T.A. Henzinger. “A survey of stochastic  $\omega$ -regular games”. *Journal of Computer and System Sciences*, vol.78 (2012), pp. 394-413.
- [10] K. Chatterjee, R. and M. Jurdziński. “On Nash equilibria in stochastic games.” *International Workshop on Computer Science Logic*. Springer Berlin Heidelberg, 2004.
- [11] A. Ehrenfeucht and J. Mycielski. “Positional strategies for mean payoff games”. *International Journal of Game Theory*, vol.8 (1979), pp. 109-113.
- [12] H. Everett. “Recursive games”. *Contributions to the Theory of Games*, vol.3 (1957), pp. 47-78.
- [13] J. Filar K. Vrieze. *Competitive Markov decision processes*. Springer Science & Business Media, 1997.
- [14] G. Hahn and G. MacGillivray, “A note on  $k$ -cop,  $l$ -robber games on graphs”. *Discrete Mathematics*, vol.306 (2006), pp.2492–2497.
- [15] Ath. Kehagias and G. Konstantinidis. “Simultaneously moving cops and robbers”. *Theoretical Computer Science*, vol. 645 (2016), pp.48-59.
- [16] R. McNaughton. “Infinite games played on finite graphs”. *Annals of Pure and Applied Logic*, vol. 65 (1993), pp. 149-184.

- [17] R. Mazala, “Infinite games.” In *Automata logics, and infinite games* (2002), pp. 23-38.
- [18] J.-F. Mertens. “Stochastic games”. *Handbook of game theory with economic applications*, vol.3 (2002), pp. 1809-1832.
- [19] R. Nowakowski and P. Winkler. “Vertex to vertex pursuit in a graph”. *Discrete Mathematics*, vol. 43 (1983), pp. 230–239.
- [20] R. Nowakowski and A. Bonato, *The Game of Cops and Robbers on Graphs*, AMS, 2011.
- [21] A. Quilliot, *Jeux et pointes fixes sur les graphes*, Ph.D. Dissertation, Universite de Paris VI, 1978.
- [22] M.J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.
- [23] L.S. Shapley. “Stochastic games”. *Proceedings of the National Academy of Sciences of the United States of America*, vol.39 (1953), pp. 1095-1100.
- [24] A. Washburn. “Deterministic graphical games”. *Journal of Mathematical Analysis and Applications*, vol.153 (1990), pp. 84-96.